SNAPSKETCH: Network Representation Approach for Anomaly Detection in Dynamic Network

- Identify denial of service attacks, port scans, and other cyber-attacks using network graphs.
- Unique approach that identifies anomalous hotspots by tracking sudden increases/decreases edges connecting to a vertex; or the sudden (dis)appearance of edges with high weight.
- The proposed SNAPSKETCH approach is fully unsupervised, has constant memory space usage, and can be used for real-time anomaly detection.

Problem Statement and Goals

**Problem Statement**
Given a graph stream $G_t = \{G_1, G_2, \ldots, G_t, \ldots\}$, our goal is to learn a graph representation function $f$ for each graph $G_t \in \mathbb{R}^{|V|}$ such that:

$$f : G_t \rightarrow v_G_t \in \mathbb{R}^d \quad \text{and} \quad d \ll |V|^2$$

and using $v_G_t$ detect whether a graph $G_t$ at any time $t$ contains an anomalous hotspot.

**Goals**
- Generate a fixed-size feature vector (SNAPSKETCH) to represent a graph in a graph stream.
- Detect DoS attack (a type of anomalous hotspot) in network traffic using a SNAPSKETCH.

Approach

- Perform node2vec [1] random walk and construct n-shingles.
- Project discriminative shingles into a $d$-dimensional projection.
- Sketch graphs using a simplified hashing of projection vector and the cost of shingles.

Results

**DoS Attack Detection Result:**
- **Smart Home IoT Traffic Data** - 95% precision and 93% recall (in 100 most severe DoS attack graphs).
- **DARPA 1998 Data** - 83% precision and 82% recall (in 100 most severe DoS attack graphs).

**Future Work:**
Integrate structural information into SNAPSKETCH for better representation.

**Acknowledgement:** Advisor Dr. William Eberle

**References:**